Prediction with expert advice: a PDE perspective on a model problem from machine learning

Robert V. Kohn
Courant Institute, NYU

Joint work with
Nadejda Drenska (now at Johns Hopkins)
Vlad Kobzar (now at Columbia) and
Zhilei Wang (now at Citadel)

Geometric PDE and Applied Analysis Seminar, OIST
November 17, 2022

Talk plan

(1) Introduction
Prediction with expert advice as a 2-person, zero-sum game

(2) A nonlinear PDE for the scaled value function
Joint work with Nadejda Drenska, JNLS 2020

(3) A PDE-informed approach to choosing strategies
Joint work with Vladimir Kobzar and Zhilei Wang, COLT 2020

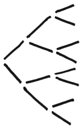
(4) Perspective
Pointers to analogous uses of PDE in other ML problems

Online machine learning

In online machine learning, information arrives sequentially and decisions must be made based on information available.

A widely-studied paradigm: predictor draws on guidance from N experts, aiming to minimize worst-case regret.

VERSION 1: Predicting a binary sequence (the stock prediction problem)



- a time series – eg a binomial stock price tree;
- a notion of gain/loss due to good/bad predictions (eg buy or sell stock);
- N experts (eg trading rules based on recent history);
- predictor's goal: do as well as the (retrospectively) best-performing expert – or at least, don't fall too far behind;
- focus on worst-case scenario (malevolent market)
- not today's focus, but PDE's are useful here too

Today's focus

VERSION 2: Predictor has no mind of his own – he just integrates the advice of many experts. So let's ignore any underlying time series.

- N experts
- predictor's action: at each time step, "choose expert to follow"

    to allow mixtures: predictor chooses a prob distrn on $\{1, \dots, N\}$ (follow expert j with prob p_j)

- adversary's action: at each time step, "choose experts' gains and losses" (eg for 3 experts, vector of gains can be $(1, -1, 1)$ or $(1, 1, -1)$ or $\dots$)

    to allow mixtures: adversary chooses a prob distrn on $\{-1, 1\}^N$

One interpretation: experts $\Leftrightarrow$ market sectors, predictor's probabilities $\Leftrightarrow$ portfolio allocations.

Prediction with expert advice as a 2-person game

Recall: predictor chooses a prob distrn (follow expert j with prob p_j); adversary chooses a prob distrn on the 2^N expert gain scenarios.

This is a 2-person, zero-sum game. The state variables are

$$x_j = j\text{th expert's gain} - \text{predictor's gain} = \text{regret wrt } j\text{th expert.}$$

The predictor's value function is:

$$U(x,t) = \text{expected final time regret, under worst-case scenario.}$$

The dynamic programming principle says (if game ends at time T):

$$\begin{aligned} U(x,t) &= \min_{\substack{\text{predictor's}\\ \text{choices}}} \max_{\substack{\text{adversary's}\\ \text{choices}}} \mathbb{E}[U(x+\Delta x, t+1)] \quad \text{for } t < T \\ U(x,T) &= \phi(x) = \max\{x_1, \cdots, x_N\} \end{aligned}$$

Talk plan

(1) Introduction

(2) A nonlinear PDE for the scaled value function
Joint work with Nadejda Drenska, JNLS 2020

(3) A PDE-informed approach to choosing strategies

(4) Perspective

Focus on long-time behavior

In ML lit, a typical question is: estimate $U(0,t)$ when $T-t$ is large, and give an easily-implemented strategy that does almost as well.

Continuum limits were designed for this; e.g. behavior of a random walk after many time steps is captured by assoc diffusion process.

So: consider scaled value function $u_\varepsilon(y,\tau) = \varepsilon U(y/\varepsilon, \tau/\varepsilon^2)$, where $\varepsilon = T^{-1/2}$ (so final time is $\tau = 1$).

- dyn prog can be written in terms of u_ε: gains become $\pm\varepsilon$ (rather than ± 1); time step is ε^2 (rather than 1).
- final-time condition is still $\phi(y) = \max_i y_i$ [other choices are possible, provided ϕ is homogenenous of degree 1].

Claim:

- There is a meaningful PDE limit; moreover, in finding it we learn about both players' optimal strategies.
- For $N = 2$, 3, and 4 experts the PDE has been solved explicitly. (So we know the optimal strategies explicitly.)

Focus on long-time behavior

In ML lit, a typical question is: estimate $U(0,t)$ when $T-t$ is large, and give an easily-implemented strategy that does almost as well.

Continuum limits were designed for this; e.g. behavior of a random walk after many time steps is captured by assoc diffusion process.

So: consider scaled value function $u_\varepsilon(y,\tau) = \varepsilon U(y/\varepsilon, \tau/\varepsilon^2)$, where $\varepsilon = T^{-1/2}$ (so final time is $\tau = 1$).

- dyn prog can be written in terms of u_ε: gains become $\pm\varepsilon$ (rather than ± 1); time step is ε^2 (rather than 1).
- final-time condition is still $\phi(y) = \max_i y_i$ [other choices are possible, provided ϕ is homogenenous of degree 1].

Claim:

- There is a meaningful PDE limit; moreover, in finding it we learn about both players' optimal strategies.
- For $N = 2$, 3, and 4 experts the PDE has been solved explicitly. (So we know the optimal strategies explicitly.)

Focus on long-time behavior

In ML lit, a typical question is: estimate $U(0,t)$ when $T-t$ is large, and give an easily-implemented strategy that does almost as well.

Continuum limits were designed for this; e.g. behavior of a random walk after many time steps is captured by assoc diffusion process.

So: consider scaled value function $u_\varepsilon(y,\tau) = \varepsilon U(y/\varepsilon, \tau/\varepsilon^2)$, where $\varepsilon = T^{-1/2}$ (so final time is $\tau = 1$).

- dyn prog can be written in terms of u_ε: gains become $\pm\varepsilon$ (rather than ± 1); time step is ε^2 (rather than 1).
- final-time condition is still $\phi(y) = \max_i y_i$ [other choices are possible, provided ϕ is homogenenous of degree 1].

Claim:

- There is a meaningful PDE limit; moreover, in finding it we learn about both players' optimal strategies.
- For $N = 2$, 3, and 4 experts the PDE has been solved explicitly. (So we know the optimal strategies explicitly.)

Finding the PDE

SIMPLE VERSION: Scaled DPP defines u_ε. We expect $u_\varepsilon \to u$. Find the PDE by replacing u_ε by u in DPP and using Taylor expansion.

FANCIER VERSION: Scaled DPP is a semi-discrete numerical scheme for the desired PDE. The simple version finds the PDE for which it's a *consistent* numerical scheme.

Recall that

predictor's choice : follow expert k with prob p_k

adversary's choice : prob distr a of experts' gains $\varepsilon(g_1, \ldots, g_N)$.

If predictor follows expert k, then scaled regret increment is

$$\Delta y = \varepsilon(g_1 - g_k, \ldots, g_N - g_k) = \varepsilon(g - g_k\vec{1}).$$

So scaled dyn prog prin is:

$$u_\varepsilon(y,\tau) = \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \max_{\substack{\text{prob distr a on} \\ g \in \{-1,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_a[u_\varepsilon(y + \varepsilon(g - g_k\vec{1}), \tau + \varepsilon^2)]$$

Finding the PDE

SIMPLE VERSION: Scaled DPP defines u_ε. We expect $u_\varepsilon \to u$. Find the PDE by replacing u_ε by u in DPP and using Taylor expansion.

FANCIER VERSION: Scaled DPP is a semi-discrete numerical scheme for the desired PDE. The simple version finds the PDE for which it's a *consistent* numerical scheme.

Recall that

$$\begin{aligned} \text{predictor's choice}: &\quad \text{follow expert } k \text{ with prob } p_k \\ \text{adversary's choice}: &\quad \text{prob distr } a \text{ of experts' gains } \varepsilon(g_1, \ldots, g_N). \end{aligned}$$

If predictor follows expert k, then scaled regret increment is

$$\Delta y = \varepsilon(g_1 - g_k, \ldots, g_N - g_k) = \varepsilon(g - g_k \vec{1}).$$

So scaled dyn prog prin is:

$$u_\varepsilon(y, \tau) = \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \max_{\substack{\text{prob distr a on} \\ g \in \{-1,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_a[u_\varepsilon(y + \varepsilon(g - g_k \vec{1}), \tau + \varepsilon^2)]$$

Finding the PDE

Substitute u_ε by u (soln of anticipated PDE) in DPP:

$$u(y,\tau) \approx \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \max_{\substack{\text{prob distr a on} \\ g \in \{-1,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_a[u(y + \varepsilon(g - g_k \vec{1}), \tau + \varepsilon^2)].$$

RHS = $u(y,\tau) + \varepsilon$[terms involving $\partial_k u$] $+ \varepsilon^2$[terms involving $\partial_{ij}^2 u$ and u_τ] $+ \ldots$

Zeroth order term $u(y,\tau)$ cancels LHS.

First order term seems dominant. But min-max of first-order term alone is a linear programming problem. Its value is 0, achieved when

> predictor's choice is $p_k = \partial_k u / (\partial_1 u + \cdots \partial_N u)$;
> adversary's choices are balanced: $\mathbb{E}_a[g_1] = \ldots = \mathbb{E}_a[g_N]$.

The predictor's strategies are fully determined but the adversary's strategies are not, so we must continue . . .

Finding the PDE

Second order term is

$$u_\tau + \max_{\mathbb{E}_a[g_j] \text{ indep of j}} \tfrac{1}{2} \sum_{k=1}^{N} p_k \mathbb{E}_a \left[\langle D^2 u\,(g - g_k \vec{1}), (g - g_k \vec{1}) \rangle \right] = 0$$

in which $p_k = \partial_k u / (\partial_1 u + \cdots \partial_N u)$. But this can be greatly simplified using that

$$u(y_1 + c, y_2 + c, \ldots, y_N + c, \tau) = u(y, \tau) + c$$

(proved by induction, provided final-time function has this property). Differentiation gives $Du \cdot \vec{1} = 1$ and $D^2 u \cdot \vec{1} = \vec{0}$, and eqn reduces to

$$\boxed{u_\tau + \max_{g \in \{-1,1\}^N} \tfrac{1}{2} \langle D^2 u\, g, g \rangle = 0}$$

Adversary's optimal strategy: given g^* that achieves the max, choose distrn a to give outcome g^* with prob $1/2$ and $-g^*$ with prob $1/2$.

Rigorous result

Theorem (Drenska-K, JNLS 2020): For the final-time function $\phi(y) = \max_i y_i$, and more generally for any final-time function ϕ st

- ϕ is nondecreasing in each y_i
- ϕ has linear growth at ∞
- $\phi(y_1 + c, \ldots, y_N + c) = \phi(y) + c$

the function u_ε (defined by the scaled dynamic program) converges as $\varepsilon \to 0$ to the unique viscosity solution of the PDE with final-time condition ϕ.

Proof: follow the Barles & Souganidis approach to convergence of numerical schemes. (Briefly: a scheme that's monotone and consistent converges to the unique viscosity solution.)

In terms of the original (unscaled) game, PDE determines scaling of the predictor's outcome after many time steps:

$$(\text{worst-case expected regret after } T \text{ steps})/\sqrt{T} \to u(0,0)$$

if for the PDE the final-time condition is imposed at $\tau = 1$.

Rigorous result

Theorem (Drenska-K, JNLS 2020): For the final-time function $\phi(y) = \max_i y_i$, and more generally for any final-time function ϕ st

- ϕ is nondecreasing in each y_i
- ϕ has linear growth at ∞
- $\phi(y_1 + c, \ldots, y_N + c) = \phi(y) + c$

the function u_ε (defined by the scaled dynamic program) converges as $\varepsilon \to 0$ to the unique viscosity solution of the PDE with final-time condition ϕ.

Proof: follow the Barles & Souganidis approach to convergence of numerical schemes. (Briefly: a scheme that's monotone and consistent converges to the unique viscosity solution.)

In terms of the original (unscaled) game, PDE determines scaling of the predictor's outcome after many time steps:

$$(\text{worst-case expected regret after } T \text{ steps})/\sqrt{T} \to u(0,0)$$

if for the PDE the final-time condition is imposed at $\tau = 1$.

Is the PDE of any use?

Well, the PDE (with final-time condition $\phi(y) = \max_i y_i$) has been solved explicitly for $N = 2, 3, 4$.

$N = 2$ goes back to T. Cover (1965). Since $D^2 u \, \vec{1} = \vec{0}$ we have $\partial_{11} u = \partial_{22} u = -\partial_{12} u$. PDE reduces to linear heat eqn

$$u_\tau + \Delta u = 0,$$

and optimal adversary chooses $g = \pm(1, -1)$ with prob $1/2$ each.

$N = 3$ soln given by Bayraktar et al (Comm PDE 2020) and Kobzar et al (COLT 2020). Formula depends on ranking of regrets: if $y_1 > y_2 > y_3$ then

$$u = \frac{y_1 + y_2 + y_3}{3} + \frac{1}{2} g(y_1 - y_2, \tau) + \frac{1}{6} g(y_1 + y_2 - 2y_3, \tau)$$

where $g(z, \tau)$ solves $g_\tau + 2g_{zz} = 0$ with final-time condition $g = |z|$. An opt'l adversary strategy is $g = \pm(1, -1, -1)$ with prob $1/2$ each.

$N = 4$ soln given by Bayraktar et al (Comm PDE 2020). An amazing formula (given again in terms of ranked regrets)!

Is the PDE of any use?

Well, the PDE (with final-time condition $\phi(y) = \max_i y_i$) has been solved explicitly for $N = 2, 3, 4$.

$N = 2$ goes back to T. Cover (1965). Since $D^2 u \vec{1} = \vec{0}$ we have $\partial_{11} u = \partial_{22} u = -\partial_{12} u$. PDE reduces to linear heat eqn

$$u_\tau + \Delta u = 0,$$

and optimal adversary chooses $g = \pm(1, -1)$ with prob $1/2$ each.

$N = 3$ soln given by Bayraktar et al (Comm PDE 2020) and Kobzar et al (COLT 2020). Formula depends on ranking of regrets: if $y_1 > y_2 > y_3$ then

$$u = \frac{y_1 + y_2 + y_3}{3} + \frac{1}{2} g(y_1 - y_2, \tau) + \frac{1}{6} g(y_1 + y_2 - 2y_3, \tau)$$

where $g(z, \tau)$ solves $g_\tau + 2g_{zz} = 0$ with final-time condition $g = |z|$. An opt'l adversary strategy is $g = \pm(1, -1, -1)$ with prob $1/2$ each.

$N = 4$ soln given by Bayraktar et al (Comm PDE 2020). An amazing formula (given again in terms of ranked regrets)!

Is the PDE of any use?

Well, the PDE (with final-time condition $\phi(y) = \max_i y_i$) has been solved explicitly for $N = 2, 3, 4$.

$N = 2$ goes back to T. Cover (1965). Since $D^2 u\,\vec{1} = \vec{0}$ we have $\partial_{11} u = \partial_{22} u = -\partial_{12} u$. PDE reduces to linear heat eqn

$$u_\tau + \Delta u = 0,$$

and optimal adversary chooses $g = \pm(1, -1)$ with prob $1/2$ each.

$N = 3$ soln given by Bayraktar et al (Comm PDE 2020) and Kobzar et al (COLT 2020). Formula depends on ranking of regrets: if $y_1 > y_2 > y_3$ then

$$u = \frac{y_1 + y_2 + y_3}{3} + \frac{1}{2} g(y_1 - y_2, \tau) + \frac{1}{6} g(y_1 + y_2 - 2y_3, \tau)$$

where $g(z, \tau)$ solves $g_\tau + 2g_{zz} = 0$ with final-time condition $g = |z|$. An opt'l adversary strategy is $g = \pm(1, -1, -1)$ with prob $1/2$ each.

$N = 4$ soln given by Bayraktar et al (Comm PDE 2020). An amazing formula (given again in terms of ranked regrets)!

Is the PDE of any use?

Most nonlinear PDE's don't have explicit solutions. (Moreover, methods used for $N \leq 4$ seem unlikely to work for larger N.)

A different idea, due to D Rokhlin (Int J Pure Appl Math 2017): since PDE has a comparison principle, an explicit subsolution or supersolution gives a bound on u. For example: if

$$w_\tau + \max_{g \in \{-1,1\}^N} \tfrac{1}{2}\langle D^2 w\, g, g\rangle \geq 0$$

with $w \leq \phi$ at the final time and $w(y + c\vec{1}, \tau) = w(y, \tau) + c$, then

$$w(y, \tau) \leq u(y, \tau).$$

A good idea, but it leaves some crucial questions:

- Where to find such w?
- We seek good strategies, not just bounds. How do they emerge? (A strategy for the predictor should give an upper bound; a strategy for the adversary should give a lower bound.)

Is the PDE of any use?

Most nonlinear PDE's don't have explicit solutions. (Moreover, methods used for $N \leq 4$ seem unlikely to work for larger N.)

A different idea, due to D Rokhlin (Int J Pure Appl Math 2017): since PDE has a comparison principle, an explicit subsolution or supersolution gives a bound on u. For example: if

$$w_\tau + \max_{g \in \{-1,1\}^N} \tfrac{1}{2} \langle D^2 w\, g, g \rangle \geq 0$$

with $w \leq \phi$ at the final time and $w(y + c\vec{1}, \tau) = w(y, \tau) + c$, then

$$w(y, \tau) \leq u(y, \tau).$$

A good idea, but it leaves some crucial questions:

- Where to find such w?
- We seek good strategies, not just bounds. How do they emerge? (A strategy for the predictor should give an upper bound; a strategy for the adversary should give a lower bound.)

Talk plan

(1) Introduction

(2) A nonlinear PDE for the scaled value function

(3) A PDE-informed approach to choosing strategies
Joint work with Vladimir Kobzar and Zhilei Wang, COLT 2020

(4) Perspective

PDE approach to adversary strategies & lower bds

Outline:

- Consider a particular strategy for the adversary (a choice of the distribution a, depending perhaps on x and t, chosen st $\pm g$ have the same probability, so $\mathbb{E}_a[g_i] = 0$ is indep of i).
- Let $V_a(x,t)$ be the predictor's best outcome against this strategy. It is characterized by a dynamic programming principle.
- Show that if $w(x,t)$ is sufficiently smooth and satisfies

$$w_t + \tfrac{1}{2}\mathbb{E}_a\langle D^2 w\, g, g\rangle \geq 0 \quad \text{for } t < T$$

$$w(x,T) \leq \max_j x_j$$

$$w(x + c\vec{1}, t) = w(x,t) + c \quad \text{for } c \in \mathbb{R}$$

then

$$V_a \geq w - (\text{error terms})$$

- Applying this, obtain the best known lower bound for large N. (The function w will solve a linear heat eqn $w_t + \kappa\Delta w = 0$.)

A PDE approach to adversary strategies & lower bds

Arguments are elementary (no viscosity theory needed)!

Given adversary strategy a, predictor's best outcome satisfies

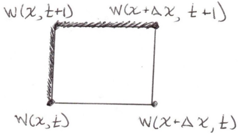
$$V_a(x,t) = \min_{\substack{\text{predictor}'\text{s} \\ \text{distrn } p}} \mathbb{E}_{p,a}[V_a(x+\Delta x, t+1)].$$

Lower bound is proved by induction (backward in time). If $V_a \geq w -$ (error terms) at time $t+1$ then

$$V_a(x,t) \geq \min_{\substack{\text{predictor}'\text{s} \\ \text{distrn } p}} \mathbb{E}_{p,a}[w(x+\Delta x, t+1)] - \text{(error terms)}.$$

Estimate RHS by considering increments in time and space:

$$\begin{aligned} w(x+\Delta x,t+1)-w(x,t) \quad &= \quad [w(x,t+1)-w(x,t)] \\ &\quad + [w(x+\Delta x,t+1)-w(x,t+1)] \end{aligned}$$



A PDE approach to adversary strategies & lower bds

$$V_a(x,t) \geq \min_p \left[\mathbb{E}_{p,a} w(x+\Delta x, t+1)\right] - (\text{error})$$

$$\begin{aligned} w(x+\Delta x, t+1) = w(x,t) &+ [w(x,t+1) - w(x,t)] \\ &+ [w(x+\Delta x, t+1) - w(x,t+1)] \end{aligned}$$

Increment in time: $w(x,t+1) - w(x,t) = w_t + \text{error}$.

Increment in space uses $\mathbb{E}[g_i] = 0$ and hypotheses on w:

$$\begin{aligned} &\textstyle\sum_k p_k \, \mathbb{E}_a[w(x + (g - g_k \vec{1}), t+1)] - w(x,t+1) \\ &\quad = \textstyle\sum_k p_k \, \mathbb{E}_a[w(x+g, t+1) - \cancel{g_k}] - w(x,t+1) \\ &\quad = \textstyle\sum_k p_k \, \mathbb{E}_a[\cancel{Dw \cdot g} + \frac{1}{2}\langle D^2 w \, g, g\rangle] + \text{error}. \end{aligned}$$

Thus $w_t + \frac{1}{2}\mathbb{E}_a\langle D^2 w \, g, g\rangle \geq 0$ implies

$$V_a \geq w - (\text{error}) \quad \text{at time } t,$$

as desired. Error terms come from Taylor expansion.

Making this concrete

Choose w to solve a linear heat eqn:

$$w_t + \kappa \Delta w = 0 \quad \text{for } t < T\text{, with } w(x,T) = \max_i x_i;$$

Since $\max_i x_i$ is homogeneous, $w(x,t) = \sqrt{T-t}\, F(x/\sqrt{T-t})$.
(Taylor expansion errors are small due to form of w.)

Large-N asymptotics: $w(0,0) = \sqrt{2\kappa T}\, \mathbb{E}_G[\max G_i]$ where each G_i is an independent standard Gaussian. From probability: $\mathbb{E}_G[\max G_i] \sim \sqrt{2 \log N}$ as $N \to \infty$, so

$$w(0,0) \sim \sqrt{4\kappa T \log N} \quad \text{as } N \to \infty.$$

Error terms are smaller, if T is large enough compared to N: their sum is at most of order $\sqrt{N \log N} + \sqrt{N} \log(T-t)$ at time t.

But: what is κ? And what is the adversary strategy a?

Making this concrete

Classic choice: if adversary chooses each g_i independently, then

$$\mathbb{E}_a\langle D^2 w\, g, g\rangle = \Delta w$$

so $w_t + \frac{1}{2} E_a\langle D^2 w\, g, g\rangle = 0$ when

$$w_t + \kappa \Delta w = 0 \quad \text{with} \quad \kappa = 1/2.$$

But one can do better! Key observation: $D^2 w$ is not just any symmetric matrix – it satisfies $D^2 \vec{1} = \vec{0}$. Using this: if adversary's choices are unif distributed over

$$S_{\text{odd}} = \{(g_1, \ldots, g_N) \,:\, g_i = \pm 1 \text{ and } \textstyle\sum_i g_i = \pm 1\} \quad \text{for } N \geq 3 \text{ odd}$$
$$S_{\text{even}} = \{(g_1, \ldots, g_N) \,:\, g_i = \pm 1 \text{ and } \textstyle\sum_i g_i = 0\} \quad \text{for } N \geq 3 \text{ even}$$

then

$$\mathbb{E}_a\langle D^2 w\, g, g\rangle = \begin{cases} (1 + \frac{1}{N})\Delta w & \text{for } N \geq 3 \text{ odd} \\ (1 + \frac{1}{N-1})\Delta w & \text{for } N \geq 3 \text{ even} \end{cases}$$

so we may take

$$\kappa = \tfrac{1}{2}(1 + \tfrac{1}{N}) \text{ for } N \text{ odd}; \quad \kappa = \tfrac{1}{2}(1 + \tfrac{1}{N-1}) \text{ for } N \text{ even.}$$

Larger κ means larger w, hence a better lower bound.

Making this concrete

Classic choice: if adversary chooses each g_i independently, then

$$\mathbb{E}_a \langle D^2 w\, g, g\rangle = \Delta w$$

so $w_t + \frac{1}{2} E_a \langle D^2 w\, g, g\rangle = 0$ when

$$w_t + \kappa \Delta w = 0 \quad \text{with} \quad \kappa = 1/2.$$

But one can do better! Key observation: $D^2 w$ is not just any symmetric matrix – it satisfies $D^2 \vec{1} = \vec{0}$. Using this: if adversary's choices are unif distributed over

$$S_{\text{odd}} = \{(g_1, \ldots, g_N) \,:\, g_i = \pm 1 \text{ and } \textstyle\sum_i g_i = \pm 1\} \quad \text{for } N \geq 3 \text{ odd}$$
$$S_{\text{even}} = \{(g_1, \ldots, g_N) \,:\, g_i = \pm 1 \text{ and } \textstyle\sum_i g_i = 0\} \quad \text{for } N \geq 3 \text{ even}$$

then

$$\mathbb{E}_a \langle D^2 w\, g, g\rangle = \begin{cases} (1 + \frac{1}{N})\Delta w & \text{for } N \geq 3 \text{ odd} \\ (1 + \frac{1}{N-1})\Delta w & \text{for } N \geq 3 \text{ even} \end{cases}$$

so we may take

$$\kappa = \tfrac{1}{2}(1 + \tfrac{1}{N}) \text{ for } N \text{ odd}; \quad \kappa = \tfrac{1}{2}(1 + \tfrac{1}{N-1}) \text{ for } N \text{ even}.$$

Larger κ means larger w, hence a better lower bound.

This adversary's algebra

Since

$$\mathbb{E}_a \langle D^2 w\, g, g \rangle = \frac{1}{|S|} \sum_{g \in S} \langle D^2 w, gg^T \rangle$$

our task is to identify $X = \frac{1}{|S|} \sum_{g \in S} gg^T$.

X has each diag element equal to 1 (obvious), and its off-diag elements are all equal (since S is permutation-invariant). So

$$X = (1 - \lambda) I + \lambda M \quad \text{where } M \text{ is the matrix of ones.}$$

Since $\langle gg^T, M \rangle = (\sum_i g_i)^2$, our choice of S gives

$$\langle X, M \rangle = 1 \text{ for } N \text{ odd}, \quad \langle X, M \rangle = 0 \text{ for } N \text{ even},$$

which implies $\lambda = -1/N$ for N odd, and $\lambda = -1/(N-1)$ for N even.

The relation $D^2 w \vec{1} = \vec{0}$ can be written as $\langle D^2 w, M \rangle = 0$. So

$$\mathbb{E}_a \langle D^2 w\, g, g \rangle = \langle D^2 w, X \rangle = (1 - \lambda) \Delta w.$$

What about upper bounds?

Outline:

- Given a particular strategy p for the predictor (depending perhaps on x and t), let $V_p(x,t)$ be the adversary's best outcome against this strategy:

$$V_p(x,t) = \max_{\substack{\text{adversary's} \\ \text{distrn } a}} \mathbb{E}_{p,a}[V_p(x + \Delta x, t+1)].$$

- Suppose $w(x,t)$ is sufficiently smooth and satisfies

$$w_t + \frac{1}{2} \max_{g_i = \pm 1} \langle D^2 w\, g, g \rangle \leq 0 \quad \text{for } t < T,$$

$$w(x,T) \geq \max\{x_1, \dots, x_N\}$$

$$w(x + c\vec{1}, t) = w(x,t) + c \quad \text{for } c \in \mathbb{R}.$$

Then choosing $p = \nabla w(x, t+1)$ at time t gives

$$V_p \leq w + (\text{error terms})$$

- With an appropriate w, this method recovers the well-known "multiplicative weights" upper bound.

What about upper bounds?

Many elements are as before: arguments are elementary; proof is by backward induction in time; errors come from Taylor expansion of w.

Main difference from lower bound is how spatial increment is handled:

$$
\begin{aligned}
\mathbb{E}_{p,a}[w(x+\Delta x, t+1) - w(x, t+1)] \\
&= \textstyle\sum_k p_k \, \mathbb{E}_a[w(x + (g - g_k\vec{1}), t+1)] - w(x, t+1) \\
&= \textstyle\sum_k p_k \, \mathbb{E}_a[w(x + g, t+1) - g_k] - w(x, t+1) \\
&= \textstyle\sum_k p_k \, \mathbb{E}_a[Dw(x, t+1) \cdot g - g_k + \frac{1}{2}\langle D^2 w(x, t+1)\, g, g\rangle] + \text{error.}
\end{aligned}
$$

If $p = Dw(x, t+1)$ then the first-order Taylor term vanishes, since

$$\sum_k p_k (Dw(x, t+1) \cdot g) = Dw(x, t+1) \cdot g = \sum_k p_k g_k.$$

So for this predictor strategy,

$$\mathbb{E}_{p,a}[w(x+\Delta x, t+1) - w(x, t+1)] \leq \frac{1}{2} \max_{g_i = \pm 1} \langle D^2 w(x, t+1)\, g, g\rangle + \text{error.}$$

Combined with the framework used for the lower bound, this shows that if $w_t + \frac{1}{2} max_{g_i=\pm 1}\langle D^2 w\, g, g\rangle \leq 0$ then

$$V_p \leq w + (\text{sum of Taylor error terms}).$$

Making this concrete

The well-known multiplicative weights upper bound can be obtained this way (D. Roklin, Int J Pure Appl Math 2017). In fact, one verifies that

$$w(x,t) = \frac{1}{\eta} \log \left(\sum_{i=1}^{N} e^{\eta x_k} \right) - \tfrac{1}{2}\eta t$$

has the required properties for any $\eta > 0$. Moreover, due to its special structure the Taylor expansion errors have signs, which are favorable; so our arguments actually show

$$V_p(x,t) \leq w(x,t)$$

with no error term. The optimal $\eta = \sqrt{(2\log N)/T}$ gives

$$\text{predictor's expected regret} \leq \sqrt{2T\log N}.$$

Upper bounds can also be obtained using other choices of w, for example soln of linear heat eqn with a well-chosen diffusion constant. (Heat eqn upper bound beats multiplicative weight bound for $N \leq 7$.)

Talk plan

(1) Introduction

(2) A nonlinear PDE for the scaled value function

(3) A PDE-informed approach to choosing strategies

(4) Perspective

- The scaled value function solves, in the limit, a nonlinear PDE.
- Its solution determines optimal strategies for both players.
- In terms of original (unscaled) game, the PDE gives us C_N st

$$\text{worst-case expected regret after } T \text{ steps} \sim C_N\sqrt{T}.$$

- While explicit solns are only available for $N = 2, 3, 4$, our PDE-based approach to strategies and bounds works for any N.
- It provides a new perspective on potential-based prediction strategies.
- It also provides an improved adversary strategy (better than advancing the experts independently).

Perspective – other ML problems

The literature on regret minimization is vast. But mostly it uses:

- specific (suboptimal) strategies for which analysis turns out to be feasible (eg multiplicative weights); and
- analysis of discrete-time dynamic programming problems – a bit like studying diffusion by considering a random walk then applying the central limit theorem.

Today's PDE-based approach is still rather new, and its impact remains to be seen. But this approach has already been used in other settings.

Starting with examples close to this talk, other variants of prediction with expert advice have been considered:

- same game, but with a random final time ("geometric stopping") (K-Drenska JNLS 2020 and K-Kobzar-Wang MSML 2020)
- same game, but with no final time ("anytime regret") (Harvey et al, FOCS 2020)

Perspective – other ML problems

PDE methods like those of this talk have also been applied in other online machine learning settings:

- the stock prediction problem with history-dependent experts (K Zhu PhD thesis 2014; K-Drenska CPAM 2022; Calder-Drenska J Fourier Anal Appl 2020 & CPAM 2022)
- drifting games (closely related to boosting) (K-Wang arXiv:2207.11405)
- the two-armed symmetric Bernoulli bandit (rather different, since predictor has incomplete information) (K-Kobzar arXiv:2202.05767)